Function of a Random Variable

X is a r.v.

X^2, $\log_e X$, e^X, $X+b$

$\sin(X)$, $\tan(X)$

Theorem: Let X be a r.v. defined on $(\Omega, \mathcal{B}, P)$. Let $g: \mathbb{R} \to \mathbb{R}$ be a measureable function. Then $Y = g(X)$ is also a r.v.

Theorem: Given a r.v. X with cdf $F_X(\cdot)$ the distribution of r.v. $Y = g(X)$, where g is measureable, can be determined.

Pf: The cdf of Y is

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y)$$

$$= P(X \in g^{-1}(-\infty, y])$$

Since g is measureable, the set $g^{-1}(-\infty, y]$ is measureable and so this term is well-defined.

Examples: Let X be a r.v. with cdf $F_X(\cdot)$.

Let $Y_1 = aX+b$, $a \neq 0$, $b \in \mathbb{R}$

$Y_2 = |X|$, $Y_3 = X^2$, $Y_4 = \log_e X$ $(X>0)$, $Y_5 = e^X$

$Y_6 = \max(X, 0)$

cdf of Y_1:

$$F_{Y_1}(y_1) = P(Y_1 \leq y_1) = P(aX+b \leq y_1)$$

$$= P\left(X \leq \frac{y_1 - b}{a}\right) \quad \text{if } a > 0$$

$$= F_X\left(\frac{y_1-b}{a}\right)$$

$$P\left(X \geq \frac{y_1-b}{a}\right) \quad \text{if } a<0$$

$$= 1 - P\left(X \leq \frac{y_1-b}{a}\right) + P\left(X = \frac{y_1-b}{a}\right)$$

$$= 1 - F_X\left(\frac{y_1-b}{a}\right) + P\left(X = \frac{y_1-b}{a}\right)$$

$$F_{Y_2}(y_2) = P\left(Y_2 \leq y_2\right) = P\left(|X| \leq y_2\right)$$

$$= P\left(-y_2 \leq X \leq y_2\right)$$

$$= P(X \le y_2) - P(X \le -y_2) + P(X = -y_2)$$

$$= \begin{cases} F_X(y_2) - F_X(-y_2) + P(X = -y_2), & \text{if } y_2 \geq 0 \\ 0, & \text{if } y_2 < 0 \end{cases}$$

$$F_{Y_3}(y_3) = P(Y_3 \le y_3) = P(X^2 \le y_3)$$

$$= P(-\sqrt{y_3} \le X \le \sqrt{y_3}), \quad y_3 \geq 0$$

$$= \begin{cases} F_X(\sqrt{y_3}) - F_X(-\sqrt{y_3}) + P(X = -\sqrt{y_3}), & y_3 \geq 0 \\ 0, & y_3 < 0 \end{cases}$$

If X is a positive r.v., let

$$Y_4 = \log_e X.$$

$$F_{Y_4}(y_4) = P(Y_4 \le y_4) = P(\log_e X \le y_4)$$

$$= P(X \le e^{y_4}) = F_X(e^{y_4})$$

* Write cdf's of Y_5, Y_6, Y_7 etc.

In case X is a discrete r.v. with pmf $p_X(x_i)$, we can consider

$$g : \{x_1, x_2, \cdots\} \to \{y_1, y_2, \cdots\}$$

$$P(Y = y_j) = P(g(X) = y_j)$$

$$= \sum_{g(x_i) = y_j} P(X = x_i) = \sum_{g(x_i) = y_j} p_X(x_i)$$

Example: $p_X(-2) = \frac{1}{5}$, $p_X(-1) = \frac{1}{6}$

$p_X(0) = \frac{1}{5}$, $p_X(1) = \frac{1}{15}$, $p_X(2) = \frac{11}{30}$

$$Y = X^2 \longrightarrow 0, 1, 4$$

$$p_Y(0) = p_X(0) = \frac{1}{5},$$

$$p_Y(1) = p_X(-1) + p_X(1) = \frac{7}{30}$$

$$p_Y(4) = p_X(-2) + p_X(2) = \frac{17}{30}$$

<u>Theorem</u>: Let X be a continuous r.v. with pdf $f_X(\cdot)$. Let $y = g(x)$ is differentiable fn.

for all x and either $g'(x) > 0 \ \forall x$ or $g'(x) < 0 \ \forall x$. Then $Y = g(X)$ is a continuous r.v. with pdf

$$f_Y(y) = f_X\left(g^{-1}(y)\right) \left| \frac{d}{dy} g^{-1}(y) \right|$$

where range of y is determined from range of x.

<u>Proof</u>: Let $g'(x) > 0 \ \forall x$. Then g is strictly increasing and so it is a one-to-one function & so g^{-1} is also

strictly increasing & $\frac{d}{dy} g^{-1}(y) > 0 \ \forall y$

The cdf of Y

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y)$$

$$= P(X \leq g^{-1}(y)) = F_X(g^{-1}(y))$$

So the pdf of Y is

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|$$

In case g is strictly decreasing $\left(\frac{d}{dy} g^{-1}(y) < 0\right)$

So $F_Y(y) = P(g(X) \leq y) = P(X \geq g^{-1}(y))$

$$= 1 - F_X(g^{-1}(y)) + P(X = g^{-1}(y)) \to 0$$

So pdf is $f_Y(y) = -f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y)$

$$= f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|$$

<u>Examples</u>: Let X have a Weibull distⁿ

$$f_X(x) = \begin{cases} 6x^2 e^{-2x^3}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

Let $y = x^3$ $\qquad Y = X^3$

$x = y^{1/3}$ $\qquad \frac{dx}{dy} = \frac{1}{3} y^{-2/3}$

So the pdf of Y is

$$f_Y(y) = 6\, y^{2/3}\, e^{-2y} \cdot \frac{1}{3}\, y^{-2/3} \quad , \quad y > 0$$

$$= \begin{cases} 2\, e^{-2y} , & y > 0 \\ 0 , & y \leq 0 \end{cases}$$

$$U = e^{-2Y} \qquad\qquad u = e^{-2y} \rightarrow y = -\frac{1}{2}\ln u$$

$$\frac{du}{dy} = -2\, e^{-2y} \quad , \qquad f_U(u) = \begin{cases} 1 , & 0 < u < 1 \\ 0 , & \text{otherwise} \end{cases}$$

Probability Integral Transform

Let X be a continuous r.v. with cdf $F_X(\cdot)$. Define r.v.

$$Y = F_X(X)$$

Then Y has a uniform dist^n on the interval $[0,1]$.

Conversely if Y has $U[0,1]$ distn &

F is a cdf of a continuous r.v,

then $X = F^{-1}(Y)$ has a cdf F.

(*) Learn some algorithms for generation of random numbers.

Thm: Let X be a cont. r.v. with pdf $f_X(x)$ and $y = g(x)$ be diff. and let $g_i^{-1}(y)$, $i = 1, \dots k$ be k inverse images. Then the pdf of Y

$$f_Y(y) = \boxed{\sum_{i=1}^{k} f(g_i^{-1}(y)) \left| \frac{d}{dy} g_i^{-1}(y) \right|}$$